


```

REFERENCE      2 (bases 1 to 1898)
AUTHORS       Mozaki,M.
TITLE         Direct Submission
JOURNAL       Submitted (25-AUG-1999) to the DDBJ/EMBL/GenBank databases, Masami
              Ohsaka National Institute for Microbial Diseases; 3-1,
              Yamadaoka, Suita-shi, Osaka 565-0871, Japan
              E-mail:mzozaki@ibkcn.osaka-u.ac.jp, Tel.:81-6-879-8338,
              Fax:81-6-879-8339)

FEATURES             Location/Qualifiers
     source           1..1898
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_line="PC-3 prostate cancer"
                     /chromosome="6"
                     /dev_stage="adult"
                     /map="6p21.2-21.3"
                     /sex="male"
                     /issue_type="prostate epithelium"
                     /note="vector:pUC18"
     gene            406..1413
                     /gene="PST"
                     /codon_start=1
                     /product="prostate ets"
                     /protein_id="BA68953.1"
                     /db_xref="GI:6721498"
                     /translation="MDSNFGSGSSPSFSLHPPTVYRTGLERAAAGAVGTERDMS
                     PPAAPADGGTSPKQVETLQVDFEATGSDGSEVDEWVGVNLTITACILNTIAPADPDGAPVQNR
                     GSDLVDPGLTLERSLEDOYSOVSVGEIVLITACILNTIAPADPDGAPVQNR
                     MTEHQRIPLPGFAEQELAEKFLCAMESEFORPRNSPLGSDVLVAHHDIIMKSAMMKER
                     TSPQALIDCAVTSMTSDVEVDSCSQPIHLMOFLKEILLRFLSHYGFRFLMLKKK
                     GFIFEDISQVARTLMGITRRKPANMYNDKLSRSIRYOYKGIIRKPIDISRLVGYFVP
                     I"

BASE COUNT        360 a   651 c   567 g   300 t

ORIGIN
alignment_scores:
    Quality:      8.00          Length:      8
    Ratio:        1.000         Gaps:        0
Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-126-945B-2_COPY_236_243 x AB031549 ..
Align seg 1/1 to: AB031549 from: 1 to: 1898

               1 ThrAspserGtUAlaspSerSer 8
               |||||||
1111 ACCGCACGCGAGGTGACTCATCA 1134

seq_name: gb_pat1.AM048159

seq_documentation_block:
LOCUS       AM048159                3317 bp    DNA                    PAT
DEFINITION sequence 12 from Patent WO0070092.
ACCESSION   AM048159
KEYWORDS    KEYWORDS
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 3317)
AUTHORS     Kaufmann,J. and Xin,H.
TITLE       Expression of ets-domain proteins in cancer
JOURNAL     Patent: WO 0070092 A, 12 23-Nov-2000;
            Chilton Corporation (US)
FEATURES             Location/Qualifiers
     source           1..3317
                     /organism="Homo sapiens"

```


gene	alignment_scores:	Length:
	Ratio:	0
	1.000	8
	Percent Similarity:	100.000
	Percent Identity:	100.000
38400	ACTGACGACGAGTGGACGAGCTTCA	38423
seq_name: gb_p6:AL157372		
1	ThirAspergillusAlaAsperserp 8	
LOCUS	AL157372	
DEFINITION	Human DNA sequence from clone RP11-375E1 on chromosome 6, complete	
ACCESSION	AL157372	
KEYWORDS	HTG.	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Ekahyoita, M;et al; Chordata; Craniata; Vertebrata; Holo-	
AUTHORS	Ekahyoita, M;et al; Chordata; Craniata; Vertebrata; Holo-	
SKUCE:JC	1 to 150660	
TITLE	Direct Submission	

COMMENTS

Requesters: clone:sequencer.ac.uk
On Nov 30, 2000 this sequence version replaced gi:11343272.
This sequence assembly data is compared from overlapping clones.
Whereas the original sequence was annotated as variations
together with the overlapping Clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone as we submit sequences with
only a small overlap as described above.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems
such as compressions and repeats, but not necessarily within known
annotated repeat sequence elements. Where the sequence is
ambiguous, there is an annotation using the 'unsure' feature key.
The following observations are used to associate primary accession
numbers given in the feature table with their source databases:
EMBL, EMBL; GenBank, F; TrEMBL, M; MIMBER, Information
on the MIMBER database.
<http://www.sanger.ac.uk/protocols/>
was generated from part of bacterial clone collection of human
chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Chr6>

Rp11-375E1 is from the library. Rp11-11.2 constructed at the Roswell Park Cancer Institute by the group of pletor de Jong. For further details see <http://dbcpac.med.buffalo.edu/>

IMPORTANT: This sequence is not the entire insert of clone Rp11-375E1. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap. The true left end of clone Rp11-375E1 is at 1 in this sequence. The true left end of clone Rp3-39032 is at 150561 in this sequence. The right end of clone Rp11-30210 is at 115044 in this sequence.

FEATURES

source

1. 150660
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/clone="Rp11-375E1"
/clone_id="Rp11-11.2"
BASE COUNT 35470 a. 38887 c. 38740 g. 37663 t
ORIGIN

alignment_scores:

Quality: 8.00 Length: 8
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

Alignment block:

us-09-126-945b-2_copy_236_243 x AL157372/rev

Align seg 1/1 to reverse of: AL157372 from: 1 to: 150660

1 ThrpSerGluValAspSerSer 8

124658 ACCGACACCGGCTGACCTCATCA 124635

seq_name: gb_baz:AF320001

seq_documentation_block:

LOCUS AF320001 7698 bp DNA BCT 27-NOV-2000
DEFINITION Acinetobacter sp. BD413 putative integral membrane protein,
putative ATP binding protein of ABC transporter, competence factor
ComA (comA), putative Htrb protein, and putative protease genes,
complete cds, and unknown genes.
ACCESSION AF320001
VERSION AF320001.1 GI:11345436

KEYWORDS

SOURCE

ORGANISM

Acinetobacter sp. BD413.
Acinetobacter sp. BD413.
Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
Acinetobacter.

1 (bases 1 to 7698)
Friedrich.A., Harteck.T. and Averhoff.B.
TITLE
Natural transformation in mesophilic and thermophilic bacteria:
identical transformation and characterization of closely related ComA-like
thermophilins HB27

2 (bases 1 to 7698)
Friedrich.A., Harteck.T. and Averhoff.B.
REFERENCE
Submitted (08-NOV-2000) Institute of Microbiology and Genetics,
University of Goettingen, Goettingen 37077,
Germany

JOURNAL

AUTHORS

TITLE

JOURNAL

FEATURES

source

1. 7698
/organism="Acinetobacter sp. BD413"
/db_xref="taxon:104611"
169. 1404
/note="ORF1"
/codon_start=1
/transl_table=11

/product="putative integral membrane protein"
/protein_id="MG34709.1"
/db_xref="GI:11345437"

/translation="MFRPSISYGLRTYRRASRNHFISFIALVSMIGTLGVAVLIT
LTVAGPRLKNNVLAIVAPOTATISPTILTMMPELAKKGVGKRVAVAPPTOLQGM
L7AQOAVAGIMVSDIDPEYKVKSTIIOHMMVAGSITMLKKEGGLVAKOMTALGIG
GANGKRLTLVEPTSPGAVRRKRVKIVIFSIGADVSDWQVYIAMDSTLRLP
LITVAENIVSSVAVYVTKRSDIALITLSDASPTITKTFWQSTVIGVGVKGA
ITVLAENIKRIONKALNRTG
1197. 2083
/note="ORF2"

/codon_start=1
/transl_table=11
/product="putative ATP binding protein of ABC transporter"
/protein_id="MG34709.1"
/db_xref="GI:11345439

2210. 4588
/gene="comA"
2210. 4588
/gene="comA"
/codon_start=1
/transl_table=11
/product="competence factor ComA"
/protein_id="MG34710.1"
/db_xref="GI:11345439"

/translation="MNGRTLECTIEPTCTIGTILFPLNAIVFYKRRKHQIGSLRALY
ILTSASFLEIGADYADSALEERLYKIMHSIDHEVYIHQSGSSDKQIKQVAVL
NIPDVWVWVPTIDSLDAKSSSLIEIGYTRITAKTPIHSLALBQADVEMFVQ
ONMKNQVIRIQLTSQVSNVSGFELHVCQOQVWQVQDLSYKRLKRTTDS
RFRPRTIYSTPROKVLILPEKMTLYCAAPGPELITLAFITLPLTLLIIRHNS
AFSLVRSASILLMDQSPFISLAAPMSISGACFLIRITQOMQADQOHOFTFRQMO
LITVALVRSOKRIFVAAIMPPTLITFIOISMPALSNLAIPIGLIIVLDIMAGCLML
FTSGLNLTLEFNLSEIFGILGDLHLPNQVYAFPLLLGLTLATVILFPL
KGVYKESVLEFPELILFSKMRPOTIISLDGSGQSIFTOQOTRILDLTGQYDE
KSTSGIGQVYVPEFSGVHOLDIPIHLHDSDHSGPLAVAILIPTMMQANSEID
LSTPSLCHSGQVETSEKTOITLSKRLMLQSTVFRNRKASVCMVLDQLQMSDKTK
LMDGMDQETETOLDQVPTQDLAVDYLHGHSDSSSTSDTLKTHPILKSNAGFEN
GTCHSAENRSGDILFLLSTTDNGHIFNFGSDGNETFTNOSTIMLQK
1421. 3534)

/note="ORF4"
/codon_start=1
/transl_table=11
/product="putative Htrb protein"
/protein_id="MG34711.1"
/db_xref="GI:11345440"

/translation="MTSSQSSKSRRLKFRISROPILSRLEKRLKAGVTLKTRT
SRITDMLKTLPTLSDQSRRTIOMAKRNLSTKFFTSIMSSQSNLSKHKES
BEKAVYPTDSGVRQIIPALKGGFTYIIPHTGSGVYKQIPILASNSLAL
IORTAKALFLVLRNENGGDFMEIPMDQAVITQDHGCTGLVHAIHLEIDRLRPH
NHSYKRFKRNPELDNLTYLEEDALNAVQVRQANRETLNATYTTDDKILSNLSL
NOT"

5668. 6684
/note="ORF5"
/codon_start=1
/transl_table=11
/product="ATP dependent protease"
/protein_id="MG34712.1"
/db_xref="GI:11345441"

/translation="MSDMPKPQNEVNSNNKTSQEMOILLEKAVLASEQRRFRNG
IFKFKIETLLEFVYLKQCTTSSSGCTSSAHLLVNNIIGTIDSSQNSVSD
TNALERAFENKSKAVALNISPQSGPVSQSDGEMIDMEIRLAKQHPKRYAVIGDMG
ASGAVYIASADEELIVPSSVSGIVGIMPYGTIGLQAKLGIDEDRTVTSQSNDDIIS
VTPVNAQREHIOVSLDNVIAHPIINAKECGRKRLKSDPAFLSPMPMGDAIOLG
VQIRQNDLTLLRELVENKRVNTIERSPESITLQKRGSELDQGLISQSVSDQLTFQD

CDS

CDS

gene

CDS

CDS

```

alignment_scores:
  Quality: 7.00
  Ratio: 1.000
  Percent Similarity: 100.000
  Percent Identity: 100.000

alignment_block:
  US-09-126-945B-2_COPY_236_434 x AF320001 ..
  Align seg 1/1 to: AF320001 from: 1 to: 7698

2 AspSerGluValAspSerSer 8
|||||
|||||
|||||
7478 GATTCAGAGATTGATTCGACC 7498

```

```

928 documentation block: 9281 bp DNA BCT 27-APR-1996
LOCUS 9281
DEFINITION Mycobacterium bovis detection region 3, 5' end.
ACCESSION U35017
VERSION 035017.1 GI:1049238
KEYWORDS
SEGMENT
SOURCE
ORGANISM
Mycobacterium bovis.
Mycobacterium bovis.
Actinobacteria; Actinobacteriia; Actinobacteridae;
Actinomycetales; Actinomycetidae; Actinomycetaceae;
1 (bases 1 to 9281)
Mahnlaars,G.G., Seb0,P.J., Hickey,M.J., Singh,D.C. and Stover,C.K.
Molecular analysis of genetic differences between Mycobacterium
bovis BCG and virulent M. bovis
J. bacteriol. 178 (5), 1274-1282 (1996)
9282
29 (bases 1 to 9281)
Mahnlaars,G.G., Seb0,P.J., Hickey,M.J., Singh,D.C. and Stover,C.K.
Direct Submission
Submitted (29-APR-1995) Mark J. Hickey, Molecular Microbiology,
Pathogenesis Corp., 201 Elliott Ave. W., Seattle, WA 98119, USA
FEATURES
source
1..9281=+"Mycobacterium bovis"
2..9281=+"Mycobacterium bovis"
misc_difference replace(1400>9281,**)
note=absent in related avirulent strain M. bovis BCG*
BASE COUNT 1706 a 3191 c 2929 g 1453 t 2 others
ORIGIN

```

```

alignment_scores:
  Quality: 7.00
  Ratio: 1.000
  Percent Similarity: 100.000
  Length: 7
  Gaps: 0
  Percent Identity: 100.000
alignment_block:
  OS-09-16-945B-2.COPY.236..243 x MBDK361/rev ..

```

```

alignment_scores: 7.00 Length: 7
                   Query: 100.000
                   Ratio: 1.000
Percent Similarity: 100.000 Percent Identity: 100.000

alignment block:
US-09-126-945B-2_COPY_236_243 x 106266/rev ..
Align seg 1/1 to reverse of: 186264 from: 1 to: 124123

```

2 AggSeqLvl=lvlaAggSeqSer 8
|||||
1459 GATTCGAGCGTAGACGCACG NC93

seq_name: qb_Ba3:MYC336

segDocumentationBlock:
DEFINITION MYC336 bacteriophage
ACCESSION J05588
VERSION 255886_1 GI:3261795

SOURCE
KEYWORDS
ORIGINIS
REFERENCE
AUTHORS

Mycobacterium tuberculosis.
Mycobacterium tuberculosis
Bacteria; Filumetes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacteriales; Mycobacteriaceae;
Mycobacterium; Mycobacterium tuberculosis complex.
Cole S.P., Broock S.V., Parshall J., Garner T., Church C.,
Harris D., Gordon S.V., Edlinger L., Tekis R.,
Tekis F., Davies K., Basham D., Brown D., Chillingworth I.,
Connor R., Bayes R., Devlin R., Fellwell T., Gentile S.,
Hamlin N., Holroyd S., Hornsby T., Jagels K., Krogh A., Mclean J.,
Moyle S., Murphy L., Oliver S., Osborne J., Quail M.A.,
Randleman M.A., Rogers J., Rutter S., Seeger K., Skelton S.,
Squires S., Squires R., Sulston J.E., Taylor K., Whitehead S., and
deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence
Nature 393 (6685), 537-544 (1998)
98295987

JOURNAL
MEDLINE
REMARK

Eratrum: [[published erratum appearing in Nature 1998 Nov
12;396(6707):190]]

REFERENCE 2 (bases 1 to 32437)
AUTHORS Parkhill,J.
TITLE Direct Submission
JOURNAL Submitted (11-JUN-1998) Submitted on behalf of the Mycobacterium tuberculosis sequencing and mapping teams, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA Unite de Genetique Moleculaire, Centre National de Recherche Scientifique, 75724 Paris Cedex 16, France
COMMENT On Jun 27, 1998 this sequence version replaced by: J117235.
Note:
Details of M. tuberculosis sequencing at the Sanger Centre are available on the World Wide Web.
(URL, http://www.sanger.ac.uk/Projects/M_tuberculosis/) CDS have been renumbered from the original contid submissions but the old gene designations are in brackets after the new gene numbers.
Gene names in brackets are from the original submissions. The gene complement in this database is based on a Hidden Markov Model of 79 genes implemented in Tigrase (Krog), supplemented with visual inspection of the genome. The complement is expected to be complete. CAUTION: In some cases we may not have predicted the correct initiation codon. Where possible we choose an initiation codon (atg, ggg, or tgg) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon.
FEATURES
source
1..32437
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
1.. 273
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
3.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
33.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
35.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
37.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
39.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
41.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
43.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
45.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
47.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
49.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
51.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
53.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
55.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
57.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
59.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
61.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
63.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
65.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
67.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
69.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
71.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
73.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
75.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
77.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
79.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
81.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
83.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
85.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
87.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
89.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
91.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
93.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
95.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
97.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
99.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
101.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
103.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
105.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
107.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
109.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
111.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
113.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
115.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
117.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
119.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
121.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
123.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
125.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
127.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
129.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
131.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
133.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
135.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
137.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
139.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
141.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
143.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
145.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
147.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
149.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
151.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
153.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
155.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
157.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
159.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
161.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
163.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
165.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
167.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
169.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
171.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
173.. 32031
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="tax

[illegible]

/evidence-not_experimental
/protein_id="BAB10289.1"
/db_xref="GI:10177214"
/translation="MASVSGCFVVTAVLAVSFLAPLAPASALVNRREYIDSCORVANKA
FCVOTLTTPPPAAAGTGLPLAFAVVGTAISHCKTAGAETAKKADIKTQREEDH
DAVQVGLASIKSALLKELPLAFAVVGSDDTNRVNGVLEKNTDITASKTLMEMTL
OMQGLDLAAGANTADADDSDNHRRA
Join(16418) : 16592,11047...17397)
/note="contains similarity to pollen-specific protein
Bmnl
gene_id:MPF21.5"
/codon_start=1
/evidence-not_experimental
/protein_id="BAB10290.1"
/db_xref="GI:10177215"
/translation="MASPSNYFLVSLAVLLEPLITASATTIDACSVTDARCA
KTLMTIPASSTQSPQAKTALMLAISTDASGPGFPAKKEPLTQPPASDAP
KLTITLANTATVAGC11PTPTANTDVACSDIATYANLVKNTDASNTVMTLME
Join(139322) : 19498,19828...20178)
/note="gene_id:MPF21.6
PII1705570
similar to unknown protein"
/codon_start=1
/evidence-not_experimental
/protein_id="BAB10291.1"
/db_xref="GI:10177216
/translation="MAGSRYVAVLVLQCLVNASATKTDICHSNARCA
KTLNAPRATSTSPSSVATVTHLQVSNANACPGACAKQTELEEDQLEI
ALISRLKASSELADYPAVDVDTCLDNATMKNLVKNTDASNTVMTLME
IPFAGKSTSIDIKT"
Join(21840) : 21989,22401...22751)
/note="contains similarity to pollen-specific protein
Bmnl
gene_id:MPF21.7"
/codon_start=1
/evidence-not_experimental
/protein_id="BAB10292.1"
/db_xref="GI:10177217"
/translation="MAVSEILPITSSGSLTKRYDALQCHVNGPCMHMLAPPA
VTAQDPLANIVINIGISVANNTGAPAPAAKKEPLLVQFSPOHEYGIQTLDM
ARQELKESPSANVDIFRCQDITMYLKDVGNDNTSKTMDMTLOHNIIDVALGA
TVALG"
Join(23846) : 24043,24366...24719)
/note="contains similarity to pollen-specific protein
Bmnl
gene_id:MPF21.8"
/codon_start=1
/evidence-not_experimental
/protein_id="BAB10293.1"
/db_xref="GI:10177218"
/translation="MTSSHCFLVSLVALLOVLLSPVSAITLTSTIKNLQMSFY
DNPFQDLQISAVPPASATGCLPLASAVTRISVLPAPAKSPAGNAKKEPSLAKQ
VTCQGVAFVSDLSMAKSELRLDRLHATDVMTCHITVYRNVLQKNTVMSKTL
Join(25238) : 25453,25665...26125)
/note="contains similarity to AP2 domain transcription
factor
gene_id:MPF21.9"
/codon_start=1
/evidence-not_experimental
/protein_id="BAB10294.1"
/db_xref="GI:10177219
/translation="MAGSRYVAVLVLQCLVNASATKTDICHSNARCA
KTLNAPRATSTSPSSVATVTHLQVSNANACPGACAKQTELEEDQLEI
ALISRLKASSELADYPAVDVDTCLDNATMKNLVKNTDASNTVMTLME
IPFAGKSTSIDIKT"
Join(29251) : 29515,29877...29998,30076...30168)
/note="gene_id:MPF21.10
unknown protein"
/codon_start=1
/evidence-not_experimental

/protein_id="BAB10295.1"
/db_xref="GI:10177220"
/translation="MGACQADVAPRVLOHPNGKEKLSVSAVYKRNKRGCSVL
LITVPLSSSSGNGGCLALRRLKRLPOTLVLGSHVRLATTTREVMGJLAKKCSL
KKSKSGDRLKEMVKAINSTRLNEDQJQMKQGEKERSIIRSQPSJOS1S8GSS5"
Join(31553) : 31591,31679...31786,31876...31932,32002...32067,
32152...32212,32296...32393)
/note="gene_id:MPF21.11
PII175792
similar to unknown protein"
/codon_start=1
/evidence-not_experimental
/protein_id="BAB10296.1"
/db_xref="GI:10177221"
/translation="MLTDCCPLCMREYNLNLERNKQRTKRPVDSNDSPEDNG
LDKRVKQGLH1QSDQNVNKGARLRYEGLGVGVYTKFEPICRLADYDVW
AKRLQVGPSTELNLEAKRQKGNKGRPSVYFAR33886,34354...34432,
Join(328058) : 33077,33469...33524,33713,33732,33864...33781
/note="contains similarity to glucose inhibited division
protein B
gene_id:MPF21.12"
/codon_start=1
/evidence-not_experimental
/protein_id="BAB10297.1"
/db_xref="GI:10177222"
/translation="MRSICSEVETKMKHCVSHLSGLRLKLPCTNSFTFLRRPH
alignment_scores:
Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent similarity: 100.000 Percent identity: 100.000
alignment_block:
US-09-126-945b-2_COPY_236_243 x AB026550/rev ..
Align seg 1/1 to reverse of: AB026550 from: 1 to: 64569
2 Aspergillus glaucus ser 8
|||||
44873 GACGAGGAAGTGCATCA 44853
seq_name: gb-pf2:AC008530
seq_documentation_block:
LOCUS AC008530 89095 bp DNA 31-MAY-2000
DEFINITION AC008530 Human chromosome 19 clone CTC-47846, complete sequence.
ACCESSION AC008530
VERSION AC008530.7 GI:8122100
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
RECORD DOB (bases 1 to 89095) Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Direct Submission
2 (bases 1 to 89095)
REFERENCE DOE Joint Genome Institute.
Direct Submission.
Submitted (03-NOV-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
Gene (bases 1 to 89095) Institute and Stanford Human Genome Center.
Direct Submission.
Submitted (31-MAY-2000) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
On May 31, 2000 this sequence version replaced gi:7711288.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.sngc.stanford.edu

CDS

CDS

CDS

CDS

CDS

CDS

CDS

Quality: Phrap Quality >=40.99.9% of Sequence:
Estimated Total Number of Errors is 0.1.

FEATURES
SOURCE
Location/Qualifiers
1..89095
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19.1606"
/clone="CNC-478M6"

BASE COUNT 27679 a 18534 c 17203 g 25679 t
ORIGIN

alignment_scores:
Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent similarity: 100.000 Percent identity: 100.000

alignment_block:
US-09-126-945b-2_COPY_236_243 x AC0086530 ..

Align seg 1/1 to: AC0086530 from: 1 to: 89095

1 THRAPSPRGUVAJASPER 7
|||||
84199 ACACATCTCGAGGTGCACCTCA 84219

seq_name: gb_hlg24:AP002018

seq_documentation_block:
LOCUS AP002018.125674 bp DNA HTG 01-JUN-2000
DEFINITION Homo sapiens clone 2058G14 map ag22.924, ***
SEQUENCING IN PROGRESS ** 19 unordered pieces.

ACCESSION AP002018.1 GI:7798578
VERSION HTG: HTGS_PHASE1
KEYWORDS Homo sapiens DNA, clone:2058G14.
SOURCE
ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
Homo sapiens
Homo sapiens clone 1 to 125674

REFERENCE
TITLE Direct Submission
AUTHORS Submitted (09-MAY-2000) Shih-Feng Tsai, National Yang-Ming
JOURNAL University, Taipei, Taiwan 11221, Republic of China
(E-mail: ympeita@ym.edu.tw, URL: http://genome.ym.edu.tw/
Tel:886-2-28267043, Fax:886-2-2826430)

COMMENT
* The contigs are represented as 100 N.
* consists of 19 contigs.
* is not known and their order in this sequence is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved. 1775: contig of 1775 bp in length
* 1776: 18751: gap of 100 bp in length
* 1876 3417: contig of 1545 bp in length
* 3418 3517: gap of 100 bp in length
* 3518 4578: contig of 1061 bp in length
* 4579 4678: gap of 100 bp in length
* 4679 6139: contig of 1461 bp in length
* 6140 6239: gap of 100 bp in length
* 6240 7913: contig of 1576 bp in length
* 7916 7913: gap of 100 bp in length
* 7916 1165: gap of 100 bp in length
* 1165 14778: contig of 3514 bp in length
* 14779 14778: gap of 100 bp in length
* 14779 16592: contig of 1814 bp in length
* 16593 16592: gap of 100 bp in length
* 16593 20750: contig of 4058 bp in length

20751 20850: gap of 100 bp in length
* 20851 24693: contig of 4042 bp in length
* 24694 24993: gap of 100 bp in length
* 24994 32048: contig of 7055 bp in length
* 32049 32148: gap of 100 bp in length
* 32149 38202: contig of 6054 bp in length
* 38203 38302: gap of 100 bp in length
* 38303 44848: contig of 6186 bp in length
* 44849 44848: gap of 100 bp in length
* 44849 54448: gap of 100 bp in length
* 54449 54448: gap of 100 bp in length
* 54449 62197: contig of 7749 bp in length
* 62198 62297: gap of 100 bp in length
* 62298 70876: contig of 8579 bp in length
* 70877 70976: gap of 100 bp in length
* 70977 83901: contig of 12925 bp in length
* 83902 84002: gap of 100 bp in length
* 84003 98894: contig of 100 bp in length
* 98895 98994: gap of 100 bp in length
* 98995 125674: contig of 2680 bp in length.
FEATURES
SOURCE
Location/Qualifiers
1..125674
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="4"
/clone="2058G14"

BASE COUNT 37414 a 25313 c 23799 g 37308 t 1840 others
ORIGIN

alignment_scores:
Quality: 7.00 Length: 7
Ratio: 1.000 Gaps: 0
Percent similarity: 100.000 Percent identity: 100.000

alignment_block:
US-09-126-945b-2_COPY_236_243 x AP002018 ..

Align seg 1/1 to: AP002018 from: 1 to: 125674

2 ASPSPRGUVAJASPER 8
|||||
103729 GACTGACATGACATCTGTC 103749

seq_name: gb_hlg24:AC016212

seq_documentation_block:
LOCUS AC016212.1160217 bp DNA HTG 16-MAR-2000
DEFINITION Homo sapiens clone Rpl1-2597, WORKING DRAFT SEQUENCE, 9 unordered
pieces.
ACCESSION AC016212.3 GI:7249075
VERSION AC016212.3 GI:7249075
KEYWORDS HTG: HTGS_PHASE1; HTGS_DRAFT.
SOURCE
ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
Homo sapiens
Homo sapiens clone Rpl1-2597
REFERENCE
TITLE Direct Submission
AUTHORS Submitted (09-MAR-2000) Balder E. Allen, N. Anderson, M.
JOURNAL University, Baltimore, Maryland 21201, USA
2 (bases 1 to 160217)
Balder E. Allen, N. Anderson, M. Balder E. Allen, N. Anderson, M.
Brown, A., Castle, A., Collinge, M., Collins, S., Collinge, A.,
Cooke, P., Dearellano, K., Dewar, K., Donnell, D., Doyle, M.,
Ferreira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D.,
Gargan, J., Gargan, S., Grant, G., Haggis, B., Heatford, A., Horton, J.,
Lechoczek, J., Lechoczek, K., Jones, K., Kam, J., Karakas, A., Klein, J.,
McKernan, K., McKernan, K., McKernan, J., McKernan, J.,
Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,

TITLE
- JOURNAL
COMMENT
All repeats were identified using RepeatMasker: Smt, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>.

FEATURES
SOURCE

McGurk, A., McKernan, K., Meldrum, J., Mollá, M., Morris, M., Morrow, J.,
Myhalocsky, J., Nahf, R., Naylor, J., Niloff, M., O'Connor, T.,
O'Donnell, P., Pavlin, B., Peterson, K., Riley, R., Roberts, D., Roy, A.,
Severy, P., Stange-Thomann, N., Stillwell, J., Stojanovic, N., Stone, C.,
Subramanian, A., Testaye, S., Tichovolsky, N., Torrealba-Miller, I.,
Vasiliiev, R., Vo, A., Wagner, A., Wheeler, J., Wu, Y., Wyman, D.,
Yasuda, M. and Zody, M.
Submitted (17-OCT-1998) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Oct 17, 1998 this sequence version replaced g1:3738104.
All repeats were identified using RepeatMasker: Smt, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>.

Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="RC1-11 human BAC library"
/map="17"
/chromosome="17"
repeat_region complement(376..440)
/rpt_family="MUTR188"
repeat_region complement(441..488)
/rpt_family="LMA1p"
repeat_region complement(499..667)
/rpt_family="Alusp"
repeat_region complement(728..922)
/rpt_family="MUTR1"
repeat_region complement(989..1107)
/rpt_family="MUTR1p"
/rpt_family="MUTR1p"
2510..2510
/rpt_family="Alusp"
2822..2845
/rpt_family="CNAAn"
3543..3701
/rpt_family="MIR"
repeat_region complement(3751..4060)
/rpt_family="Alusp"
4303..4364
/rpt_family="L2"
4451..4610
/rpt_family="L2"
repeat_region complement(5380..5541)
/rpt_family="MER5A"
6403..6488
/rpt_family="MER5A"
repeat_region complement(6419..6499)
/rpt_family="MER5B"
7137..7211
/rpt_family="L2"
repeat_region complement(7696..7734)
/rpt_family="MER5B"
repeat_region complement(7735..8051)
/rpt_family="Alusp"
repeat_region complement(8052..8160)
/rpt_family="MER5B8216"
repeat_region complement(8161..8183)
/rpt_family="L1M5"
repeat_region complement(8217..8517)
/rpt_family="Alusp"
repeat_region complement(8518..9493)
/rpt_family="L1M5"
repeat_region complement(9500..9542)
/rpt_family="MER5B"
11520..11577
/rpt_family="MIR"
11962..12039
/rpt_family="MIR"
12069..12934
/rpt_family="L2"
repeat_region complement(13277..13326)

repeat_region /rpt_family="CNAAn"
1448..1448
/rpt_family="L2"
14961..15091
/rpt_family="MER56"
15575..15982
/rpt_family="L1R16B"
16458..16738
/rpt_family="L1"
repeat_region /rpt_family="L12"
17121
/rpt_family="MIR"
repeat_region complement(17298..17441)
/rpt_family="purine-rich"
18010..18297
/rpt_family="Alusp"
18569..18610
/rpt_family="MDEL1"
repeat_region complement(18967..19251)
/rpt_family="MER7C"
19252..19585
/rpt_family="Alusp"
repeat_region complement(19557..19860)
/rpt_family="MER7C"
20913..21017
/rpt_family="MIR"
21034..21261
/rpt_family="L2"
21557..21569
/rpt_family="L2"
21588..21888
/rpt_family="Alusp"
repeat_region complement(22305..22557)
/rpt_family="Alusp"
repeat_region complement(22820..23126)
/rpt_family="Alusp"
24275..24536
/rpt_family="MIR"
24823..25310
/rpt_family="L2"
25217..25489
/rpt_family="Alusp"
repeat_region complement(25430..25509)
/rpt_family="L2"
repeat_region complement(26510..26597)
/rpt_family="MIR"
28125..28482
/rpt_family="L1MEL1"
repeat_region complement(28532..29819)
/rpt_family="Aluv"
repeat_region complement(31605..33787)
/rpt_family="L2"
34769
/rpt_family="MIR"
repeat_region complement(35265..35311)
/rpt_family="CNAAn"
repeat_region complement(35478..35498)
/rpt_family="CNAAn"
35832..35929
/rpt_family="MER5A"
repeat_region complement(36458..36712)
/rpt_family="Alusp"
36923..36972
/rpt_family="Alusp"
repeat_region /rpt_family="MIR"
36973..37271
/rpt_family="Aluv"
repeat_region 37272..37477
/rpt_family="MIR"
repeat_region complement(37478..37524)
/rpt_family="CNAAn"
37635..37821
/rpt_family="MIR"

```
repeat_region 37895..38191  
/rpl_family-"L1"  
repeat_region complement(38192..38501)  
/rpl_family-"Alusg"  
repeat_region 38502..38768  
/rpl_family-"L1"  
repeat_region 38768..39400  
/rpl_family-"11NC/D"  
repeat_region complement(39478..39766)  
/rpl_family-"Aluo"  
repeat_region 39767..39801  
/rpl_family-"AR-rich"  
repeat_region complement(42375..42461)
```

```
alignment_scores:  
Quality: 7.00 Length: 7  
Ratio: 1.000 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000
```

```
alignment_block:  
US-09-126-945b-2_COPY_236_243 x AC005747 ..
```

```
Align seg 1/1 to: AC005747 from: 1 to: 166701
```

```
2 AppSerGIuVaIAspSerSer 8  
|||||  
136917 GACAGTGAGGTGGACACAC 136937
```

THIS PAGE BLANK (U8PTD)